

EXHIBIT 1



PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

Books

Entrez Nucleotide

My NCBI
[Sign In] [Register]

Search Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full)

Show 5

Send to

Projection on the feature (#6841)

Show whole Sequence

Features: +

Refresh

☐ 1: BA000043. Reports *Geobacillus kaust...*[gi:56378377]

Links

Features Sequence

LOCUS BA000043 1137 bp DNA linear BCT 04-DEC-2004
 DEFINITION *Geobacillus kaustophilus* HTA426 DNA, complete genome.
 ACCESSION BA000043 REGION: 1613..2749
 VERSION BA000043.1 GI:56378377
 KEYWORDS .
 SOURCE *Geobacillus kaustophilus* HTA426
 ORGANISM *Geobacillus kaustophilus* HTA426
 Bacteria; Firmicutes; Bacillales; Bacillaceae; *Geobacillus*.
 REFERENCE 1
 AUTHORS Takami,H., Takaki,Y., Chee,G.J., Nishi,S., Shimamura,S., Suzuki,H.,
 Matsui,S. and Uchiyama,I.
 TITLE Thermoadaptation trait revealed by the genome sequence of
 thermophilic *Geobacillus kaustophilus*
 JOURNAL (er) Nucleic Acids Res. 32 (21), 6292-6303 (2004)
 PUBMED 15576355
 REFERENCE 2 (bases 1 to 1137)
 AUTHORS Takami,H., Takaki,Y. and Chee,G.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUN-2003) Hideto Takami, Japan Marine Science and
 Technology Center, Microbial Genome Analysis Research Group; 2-15
 Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan
 (E-mail:takamih@jamstec.go.jp,
 URL:http://www.jamstec.go.jp/jamstec-e/bic/exbase.html,
 Tel:81-46-867-9643, Fax:81-46-867-9645)
 FEATURES Location/Qualifiers
 source 1..1137
 /organism="Geobacillus kaustophilus HTA426"
 /mol_type="genomic DNA"
 /strain="HTA426"
 /isolation_source="isolated from the deepest Ocean"
 /db_xref="taxon:235909"
 /note="thermophile"
 gene 1..1137
 /gene="dnaN"
 /locus_tag="GK0002"
 CDS 1..1137
 /gene="dnaN"
 /locus_tag="GK0002"
 /EC_number="2.7.7.7"
 /codon_start=1
 /transl_table=11
 /product="DNA-directed DNA polymerase III beta subunit"
 /protein_id="BAD74287.1"
 /db_xref="GI:56378379"
 /translation="MNISIDREALAKSVQDVMKAVSTRRTTIPILTGIKLTATASGVTL
 TGSDSDISIESFIPLEKEGKLLVDVKRPGSIVLQARFFSEIVKKLPQQTVEIETEDNF
 LTIIRSGHSEFRLNGLNADEYPRLPQIEENVFQIPADLLKTVIRQTVFAVSTSETRP

ILTGNNWKVERGELVCTATDSHRLAMRKVKIESENEVSYNVVI PGKSLNELSKILDDG
NHPVDIVMTANQVLFKAERHLLFFSRLLDGNYPETARLIPTESKTTMIVNTKEFLQAIID
RASLLAREGRNNVVKLTTLPGGMLEISSISPEIGKVTEQLQTESLEGEELNISFSAKY
MMDALRALDGTDTIQISFTGAMRPFLLRPLHTDSMLQLILPVRTY"

ORIGIN

```
1  gtgaacattt  ccattgaccg  cgaagcggtg  gccaaagagcg  tccaagacgt  gatgaaagcc
61  gtatogacaa  gaacgacgat  tccgattttt  aocgggaatca  aactaacagc  aacgcgctcg
121 ggagtaacgc  tgaccggaag  cgattccgac  atttcaattg  aatcgtttat  tccgottgaa
181 aaagaaggca  agttgctcgt  tgacgtgaaa  agaccgggga  gcacgtact  gcaggcgcg
241 tttttctctg  aaatcgtgaa  aaaaactgcc  caacaaacgg  tggaaatcga  aacggaagac
301 aactttttga  cgatcatccg  ctccggggc  tcagaattcc  gccccaatgg  gctaaacgcc
361 gacgaatatc  cgcgcctgcc  gcaaattgaa  gaagaaaacg  tgtttcaaat  cccggctgat
421 ttattgaaaa  ccgtgattcg  gcaaacggtg  ttccgcgttt  ctacatcgga  aacgcgcccc
481 atcttgacag  gtgtcaactg  gaaagttgaa  catggcgagc  ttgtctgcac  agcgaccgac
541 agtcacatcg  tagccatcgc  caaagtgaag  attgagtcgg  aaaatgaagt  atcatacaac
601 gtgcgcaccc  ctggaaaaag  tcttaattgag  ctccgcaaaa  ttttgatga  cggcaaccac
661 ccggtggaca  tcgtcatgac  agccaatcaa  gtgctattta  agcccgagca  ccttctcttc
721 ttttcccgcc  tgcttgacgg  caactatccg  gagacggccc  gottgattcc  aacagaaagc
781 aaaacgacca  tgatcgtcaa  tacaaaagag  tttctgcagg  caatcgaccg  agcgtccttg
841 cttgctcgag  aagggaaggaa  caacgttgtg  aaactaacga  cgcttctctg  aggaatgctc
901 gaaattttct  cgattttctc  ggagatcggg  aaagtgaagg  agcagctgca  aacggagtct
961 cttgaagggg  aagagttgaa  catttcgttc  agcgcgaaat  atatgatgga  cgcgttgagg
1021 gcgcttgatg  gaacagacat  tcaaatcagc  ttcactgggg  ccattgcggc  gttcctgttg
1081 cgcgcgcttc  ataccgattc  gatgcttcag  ctcatcttgc  cggtgagaac  atattga
```

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Jul 24 2006 17:22:14



PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

Books

Nucleotide

My NCBI

[Sign In] [Register]

Search Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full) Show 5 Send to

Projection on the feature (#5043)

Show whole Sequence

Features:

+

Refresh

1: NC_006274. Reports *Bacillus cereus* E...[gi:52140164]

Links

Comment Features Sequence

LOCUS NC_006274 1140 bp DNA linear BCT 03-DEC-2005
 DEFINITION *Bacillus cereus* E33L, complete genome.
 ACCESSION NC_006274 REGION: 1927..3066
 VERSION NC_006274.1 GI:52140164
 PROJECT GenomeProject:12468
 KEYWORDS .
 SOURCE *Bacillus cereus* E33L
 ORGANISM *Bacillus cereus* E33L
 Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*; *Bacillus cereus* group.
 REFERENCE 1 (bases 1 to 1140)
 AUTHORS Brettin,T.S., Bruce,D., Challacombe,J.F., Gilna,P., Han,C., Hill,K., Hitchcock,P., Jackson,P., Keim,P., Longmire,J., Lucas,S., Okinaka,R., Richardson,P., Rubin,E. and Tice,H.
 TITLE Complete genome sequence of *Bacillus cereus* E33L
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1140)
 CONSRM NCBi Genome Project
 TITLE Direct Submission
 JOURNAL Submitted (16-SEP-2004) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA
 REFERENCE 3 (bases 1 to 1140)
 AUTHORS Brettin,T.S., Bruce,D., Challacombe,J.F., Gilna,P., Han,C., Hill,K., Hitchcock,P., Jackson,P., Keim,P., Longmire,J., Lucas,S., Okinaka,R., Richardson,P., Rubin,E. and Tice,H.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-2004) Joint Genome Institute, Department of Energy, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT PROVISIONAL REFSEQ: This record has not yet been subject to final NCBI review. The reference sequence was derived from CP000001.
 COMPLETENESS: full length.
 FEATURES
 source
 1..1140
 /organism="*Bacillus cereus* E33L"
 /mol_type="genomic DNA"
 /strain="E33L"
 /isolation_source="dead zebra carcass"
 /db_xref="taxon:268681"
 /country="Namibia: Etosha National Park"
 /note="original swab taken from carcass of dead zebra in April 1996; this organism was isolated in 2000"
 gene
 1..1140
 /gene="dnaN"
 /locus_tag="BCZK0002"
 /db_xref="GeneID:3024751"
 CDS
 1..1140
 /gene="dnaN"

```
/locus_tag="BCZK0002"  
/EC_number="2.7.7.7"  
/inference="non-experimental evidence, no additional  
details recorded"  
/note="binds the polymerase to DNA and acts as a sliding  
clamp"  
/codon_start=1  
/transl_table=11  
/product="DNA polymerase III subunit beta"  
/protein_id="YP_081621.1"  
/db_xref="GI:52145207"  
/db_xref="GeneID:3024751"  
/translation="MRFTTIQKDYLVRSVQDVMKAVSSRTTIPILTGIKVVATEEGVTL  
TGSDADISIESFIPVEEDGKEIVEVKQSGSIVLQAKYFSEIVKKLPKETVEISVENHL  
MTKITSGKSEFNINGLDSA EYPLLPQIEHHVFKIPTDLLKHMIRQTVFAVSTSETRP  
ILTG VNWKVYNSELTCTATDSHRLALRRAKIEGIADEFQANVVIPGKSLNELSKILDE  
SEEMVDIVITEYQVLFRTKHLFFSRLLLEGNYPDTRLIPAESKTDIFVNTKEFLQAI  
DRASLLARDGRNNVVKLSTLEQAMLEISSNSPEIGKVVEEVQCEKVDGEELKISFSAK  
YMMDALKALDSTEIKISFTGAMRPFILRTVNDESIQLILPVRTY"
```

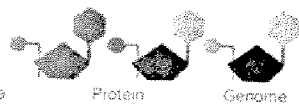
ORIGIN

```
1 atgcgctttta caattcaaaa agactatcctt gtaagaagtg tacaagatgt aatgaaggct  
61 gtttcttctc gtacaacaat tcgatccctt acaggaatta aggttgtagc tacggaagaa  
121 ggagttacat taacaggaag cgatgctgat atttcgattg aatcttttat cccagttgaa  
181 gaggatggaa aagaaatcgt agaagtaaaa caatcaggaa gtattgtttt acaggcgaaa  
241 tatttttagtg aaattgtaaa aaaattacct aaagaaactg tagaaatttc tgtggaaaat  
301 catttgatga caaaaataac ttctggaaaa tcagagttta atttaaatgg ttttagattct  
361 gcagaatata cgttgttacc acaaattgaa gaacatcatg tttttaaat cccaacggat  
421 ttacttaaac atatgatcag acaaactgta tttgcagttt ccacttctga aacaagaccg  
481 atcttgacag gtgtaaactg gaaggtatat aacagcgaac taacttgcac tgotacagat  
541 agccacaggt tagctcttcg aaaagcaaaa atcgaaggca ttgcagatga attccaagca  
601 aatgttggtta tcccagggaa aagcttaaat gaattaagca aaattttaga tgagtctgaa  
661 gaaatggtag atatcgttat tacggagtat caagtattat tccgtacaaa acatttatta  
721 ttcttctcaa gattgttaga aggaaattat cctgatacaa ctgccttaat tccagcggag  
781 agtaaaacag atatttttgt aaatacaaaa gaatttttac aagcaattga tcgtgcatcc  
841 ctattagcaa gagatggtcg taataatggt gtaaaattat caactttaga gcaggcaatg  
901 ctagaaattt cttcaaatto accagaaato gggaaagtag tagaagaagt tcaatgtgaa  
961 aaagtagatg gagaagagtt aaaaatatct tttagtgcac aatatatgat ggatgcacta  
1021 aaggcattag atagtactga gattaagatt agctttactg gagcaatgag accattctta  
1081 attcgtagcg taaatgatga atccattatt caattaattt taccggttcg tacttactaa
```

//

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

Jul 24 2006 17:22:14



My NCBI
[Sign In] [Register]
Books

Search Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full) Show 5 Send to

Projection on the feature (#1) Show whole Sequence

Features:

+

Refresh

☐ 1: NC_005957. Reports *Bacillus thuringi...*[gi:49476684]

Links

Comment Features Sequence

LOCUS NC_005957 1140 bp DNA linear BCT 03-APR-2006
 DEFINITION *Bacillus thuringiensis* serovar konkukian str. 97-27, complete genome.
 ACCESSION NC_005957 REGION: 1928..3067
 VERSION NC_005957.1 GI:49476684
 PROJECT GenomeProject:10877
 KEYWORDS
 SOURCE *Bacillus thuringiensis* serovar konkukian str. 97-27
 ORGANISM *Bacillus thuringiensis* serovar konkukian str. 97-27
 Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*; *Bacillus cereus* group.
 REFERENCE 1 (bases 1 to 1140)
 AUTHORS Brettn,T.S., Bruce,D., Challacombe,J.F., Gilna,P., Han,C., Hill,K., Hitchcock,P., Jackson,P., Keim,P., Longmire,J., Lucas,S., Okinaka,R., Richardson,P., Rubin,E. and Tice,H.
 TITLE Complete genome sequence of *Bacillus thuringiensis* 97-27
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1140)
 CONSRTM NCBI Genome Project
 TITLE Direct Submission
 JOURNAL Submitted (11-SEP-2004) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA
 REFERENCE 3 (bases 1 to 1140)
 AUTHORS Brettn,T.S., Bruce,D., Challacombe,J.F., Gilna,P., Han,C., Hill,K., Hitchcock,P., Jackson,P., Keim,P., Longmire,J., Lucas,S., Okinaka,R., Richardson,P., Rubin,E. and Tice,H.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2004) Joint Genome Institute, Department of Energy, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT PROVISIONAL REFSEQ: This record has not yet been subject to final NCBI review. The reference sequence was derived from AE017355. *Bacillus thuringiensis* 97-27 (subsp. konkukian (serotype H34)) was originally isolated from a case of severe human tissue necrosis (*Bacillus thuringiensis* subsp. konkukian (serotype H34) superinfection: Case report and experimental evidence of pathogenicity in immunosuppressed mice. Hernandez, E, Ramisse, F, Ducoureaux, J-P, Cruel, T, and Cavallo, J-D. J Clin Microbiol 1998 36(7):2138-2139). *B. thuringiensis* is indigenous to many habitats worldwide; these include soil, insects, deciduous and coniferous leaves (Prediction of insecticidal activity of *Bacillus thuringiensis* strains by polymerase chain reaction product profiles. Carozzi, NB, Kramer, VC, Warren, GW, Evola, S, and Koziel, MG. Appl Environ Microbiol. 1991 57(11):3057-61). *B. thuringiensis* is an insect pathogen that is widely used as a biopesticide in commercial agriculture. Infection of humans is unusual. The apparent pathogenic properties of *B. thuringiensis* 97-27 are very unusual for *B. thuringiensis*; unlike most *B.*

thuringiensis isolates, this isolate is very closely related to *B. anthracis* based on phylogenetic analysis (Fluorescent amplified fragment length polymorphism analysis of *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* isolates. Hill, KK, Ticknor, LO, Okinaka, RT, Asay, M, Blair, H, Bliss, KA, Laker, M, Pardington, PE, Richardson, AP, Tonks, M, Beecher, DJ, Kemp, JD, Kolsto, A-B, Wong, ACL, Keim, P, and Jackson, PJ. Appl Environ Microbiol 70(2):1068-1080. Plasmid and fosmid libraries were prepared at the Joint Genome Institute in Los Alamos (JGI-LANL), NM. Shotgun sequencing was performed at the JGI Production Genomics Facility (JGI-PGF) in Walnut Creek, CA to a coverage of 24x. Finishing was performed at JGI-LANL starting with 83 contigs and 16 scaffolds. Repetitive regions were identified, assembled and finished by manually checking paired reads close to each repeat in the assembly with consed and then making a subassembly for each repetitive region. Fifty five gaps were closed with primer walks and 16 by PCR. Gene predictions were obtained using Glimmer and tRNAs were identified using tRNAscan-SE. Basic analysis of the gene predictions was performed by comparing coding sequences against the PFam, BLOCKS and Prodom databases. Gene definitions and functional classes were added manually by a team of annotators at JGI-LANL, using BLAST results in addition to information from the basic analysis. A total of 5540 features have been annotated on the sequence record.

COMPLETENESS: full length.

```

FEATURES             Location/Qualifiers
     source            1..1140
                        /organism="Bacillus thuringiensis serovar konkukian str.
                        97-27"
                        /mol_type="genomic DNA"
                        /strain="97-27"
                        /serovar="konkukian"
                        /db_xref="taxon:281309"
     gene              1..1140
                        /gene="dnaN"
                        /locus_tag="BT9727_0002"
                        /db_xref="GeneID:2857308"
     CDS               1..1140
                        /gene="dnaN"
                        /locus_tag="BT9727_0002"
                        /EC_number="2.7.7.7"
                        /inference="non-experimental evidence, no additional
                        details recorded"
                        /note="binds the polymerase to DNA and acts as a sliding
                        clamp"
                        /codon_start=1
                        /transl_table=11
                        /product="DNA polymerase III subunit beta"
                        /protein_id="YP_034362.1"
                        /db_xref="GI:49476685"
                        /db_xref="GeneID:2857308"
                        /translation="MRESIQKDYLVRSVQDVMKAVSSRTTIPILTGIKVVATEEGVTL
                        TGSDADISIESFIPVEEDGKEIVEVKQSGSIVLQAKYFSEIVKKLPKETVEISVENHL
                        MTKITSGKSEFNGLDLSAEYPLLPQIEHHVFKIPTDLLKHMIRQTVFAVSTSETRP
                        ILTGWNWKVYNSELTCTATDSHRLALRKAKIEGIVDEFQANVVIPKSLNELSKILDE
                        SEEMVDIVITEYQVLFRTKHLLEFSRLLEGNYPDTRLIPAESKTDIFVNTKEFLQAI
                        DRASLLARDGRNNVVKLSTLEQAMLEISSNSPEIGKVVEEVQCEKVDGEELKISFSAK
                        YMDALKALDSTEIKISFTGAMRPFLLIRTVNDESIIQLILPVRTY"

ORIGIN
1  atgcggttttt caattcaaaa agactatctt gtaagaagtg tacaagatgt aatgaaggct
61  gttttcttctc gtacaacaat tccgatcctt acaggaatta aagttgtcgc tacggaagaa
121 ggagttacat taacaggaag cgatgctgat atttcgattg aatcttttat ccaggttgaa
181 gaggatggaa aagaaattgt agaagtaaaa caatcaggaa gtattgtttt acaggttaaa

```

241 tatttttagtg aaattgtaaa aaaattgccc aaagaaactg tagaaatttc tgtcgaaaaat
301 cattttaatga caaaaataac ttctgggaaa tcagaattta atttaaatgg tttagattct
361 gcagaatata cattgttacc acaaattgaa gaacatcatg tttttaagat tccaacagat
421 ttacttaaac atatgatcag acaaaactga tttgcagtct ccacttctga aacaagacca
481 atcttgacag gtgtaaactg gaaggtatat aacagcgaac taacttgtat tgctacagat
541 agtcacaggt tagctcttcg aaaagcaaaa attgaaggta ttgtagatga attccaggca
601 aatgtttgta ttccggggaa aagctttaat gaattaagca aaattctaga tgagtctgaa
661 gaaatggtag atatcgttat taoggagtat caagtattat tcogtacaaa acatttatta
721 ttcttctcaa gattgtttaga aggaattat cctgatacaa ctcgattaat tcctgcagag
781 agtaaaacag atatttttgt aaatacaaaa gaatttttac aagcaattga tcgtgcactc
841 ctgttagcta gagatggtcg taataatggt gtgaaattat caactttaga gcaggcaatg
901 ctagaaattt cttcgaattc accagaaatc gggaaagtag tagaagaagt tcaatgtgaa
961 aaagtagatg gagaagattt aaaaatatct tttagtcaa aatatatgat ggatgcacta
1021 aaggcattag atagtactga aattaagatt agctttactg gagcaatgag accattctta
1081 attcgtacgg taaatgatga atccattatt caattaattt taocgggttcg taactactaa

//

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

Jul 24 2006 17:22:14



Search Protein

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenPept

Show 20

Send to

Range: from begin to end

Features: ☒ CDD

+

Refresh

☐ 1: EAR74115. Reports DNA polymerase III...[gi:89154071]BLink, Conserved
Domains, Links

Comment Features Sequence

LOCUS EAR74115 381 aa linear BCT 03-MAR-2006
DEFINITION DNA polymerase III, beta chain [Bacillus weihenstephanensis KBAB4].
ACCESSION EAR74115
VERSION EAR74115.1 GI:89154071
DBSOURCE accession [AA0Y01000054.1](#)
KEYWORDS .
SOURCE Bacillus weihenstephanensis KBAB4
ORGANISM [Bacillus weihenstephanensis KBAB4](#)
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE 1 (residues 1 to 381)
AUTHORS Lapidus,A., Goltsman,E., Copeland,A., Lucas,S., Barry,K.,
Detter,J.C., Glavina del Rio,T., Hammon,N., Israni,S., Dalin,E.,
Tice,H., Bruce,D., Pitluck,S., Richardson,P., Auger,S.,
Galleron,N., Sanchis,V., Gohar,M., Broussole,V., Brillard,J.,
Guinebretiere,M.-H., Ehrlich,S.D., Lereclus,D., Aymerich,S.,
Nguen-the,C. and Sorokin,A.
CONSRTM US DOE Joint Genome Institute (JGI-PGF)
TITLE Sequencing of the draft genome and assembly of Bacillus
weihenstephanensis KBAB4
JOURNAL Unpublished

REFERENCE 2 (residues 1 to 381)
AUTHORS Larimer,F. and Land,M.
CONSRTM US DOE Joint Genome Institute (JGI-ORNL)
TITLE Annotation of the draft genome assembly of Bacillus
weihenstephanensis KBAB4
JOURNAL Unpublished

REFERENCE 3 (residues 1 to 381)
AUTHORS Lapidus,A., Goltsman,E., Copeland,A., Lucas,S., Barry,K.,
Detter,J.C., Glavina del Rio,T., Hammon,N., Israni,S., Dalin,E.,
Tice,H., Bruce,D., Pitluck,S., Richardson,P., Auger,S.,
Galleron,N., Sanchis,V., Gohar,M., Broussole,V., Brillard,J.,
Guinebretiere,M.-H., Ehrlich,S.D., Lereclus,D., Aymerich,S.,
Nguen-the,C. and Sorokin,A.
CONSRTM US DOE Joint Genome Institute (JGI-PGF)
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2006) US DOE Joint Genome Institute, 2800
Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA

COMMENT URL -- <http://www.jgi.doe.gov>
Contact: Paul Richardson (microbes@cuba.jgi-psf.org)
Draft sequencing done at US DOE Joint Genome Institute
Source DNA and bacteria available from Alexei Sorokin
(alexei.sorokine@jouy.inra.fr)
The JGI and collaborators endorse the principles for the
distribution and use of large scale sequencing data adopted by the
larger genome sequencing community and urge users of this data to

follow them. It is our intention to publish the work of this project in a timely fashion and we welcome collaborative interaction on the project and analysis.

(<http://www.genome.gov/page.cfm?pageID=10506376>)

Notes:

Bacillus weihenstephanensis KBAB4 was originally isolated from forest soil near Versailles, France, and was originally identified as *Bacillus cereus* (Vilas-Boas et al, Appl Env Microbiol, 2002, 68, 1414). More detailed phylogenetic analysis, using MLST, of the Versailles Collection of *B. cereus* and *B. thuringiensis* strains revealed that the strain grows at low temperature (60C) and clusters with many other strains able to grow at low temperatures, including the independently isolated and characterized strains WSBC10204 and WSBC10206 (Sorokin et al, Appl Env Microbiol, 2006, 72, 1569). The latter two strains are the type strains of a newly recognized species in the *B. cereus* group able to grow in cold and having the species name *Bacillus weihenstephanensis* (Lechner et al, Int. J. Syst. Bacteriol, 1998, 48, 1373). Since, by MLST, KBAB4 is very closely related to WSBC10204 and represents many other psychrotrophic strains, it was assigned a species name *Bacillus weihenstephanensis* and the strain KBAB4 should also be considered as a type representative of this species.

Method: conceptual translation.

```

FEATURES             Location/Qualifiers
    source             1..381
                       /organism="Bacillus weihenstephanensis KBAB4"
                       /strain="KBAB4"
                       /db_xref="taxon:315730"
    Protein            1..381
                       /product="DNA polymerase III, beta chain"
                       /EC_number="2.7.7.7"
    Region             1..379
                       /region_name="beta_clamp"
                       /note="Beta clamp domain; cd00140"
                       /db_xref="CDD:29053"
    Region             1..379
                       /region_name="beta_clamp"
                       /note="Beta clamp domain; cd00140"
                       /db_xref="CDD:29053"
    Region             17..376
                       /region_name="POL3Bc"
                       /note="DNA polymerase III beta subunit; smart00480"
                       /db_xref="CDD:47786"
    Region             17..376
                       /region_name="POL3Bc"
                       /note="DNA polymerase III beta subunit; smart00480"
                       /db_xref="CDD:47786"
    CDS                1..381
                       /locus_tag="BcerKBAB4DRAFT_2067"
                       /coded_by="AAOY01000054.1:20757..21902"
                       /transl_table=11
                       /db_xref="InterPro:IPR001001"

ORIGIN
    1 mrftiqkdyl vrsvqdvkma vssrttipil tgikvvatee gvtltgsdad isiesfipve
    61 dagkeiveie qsgsilgkak yfseivkkip ketveisven hfmtkiksgk sefnlnglda
   121 aeypllpqie ehvfkiptd lkhmirqtv favsssetrp iltgvnwkvv nseltciatd
   181 shrlalrkak iegyniadev qanvvipgks lselskilde seemvdivit eyqvlfrtkh
   241 llffsrllleg nypdtttrlip aesktdivfn tkeflgaidr asllardgrn nvvklstleq
   301 qmleissnap eigkvveevq cenvdgeelk isfsakymmd alkaldstei kvsftgamrp
   361 flirtvndds iiqlilpvrt y
//

```

Disclaimer | Write to the Help Desk



PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

My NCBI

[Sign In] [Register]

Books

Search Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full) Show 5 Send to

Projection on the feature (#2) Show whole Sequence

Features: +

Refresh

☐ 1: NC_000964. Reports *Bacillus subtilis*...[gi:50812173]

Links

Comment Features Sequence

LOCUS NC_000964 1137 bp DNA linear BCT 02-DEC-2005

DEFINITION *Bacillus subtilis* subsp. *subtilis* str. 168, complete genome.

ACCESSION NC_000964 REGION: 1939..3075

VERSION NC_000964.2 GI:50812173

PROJECT GenomeProject:76

KEYWORDS complete genome.

SOURCE *Bacillus subtilis* subsp. *subtilis* str. 168

ORGANISM *Bacillus subtilis* subsp. *subtilis* str. 168
Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.

REFERENCE 1 (bases 1 to 1137)

AUTHORS Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borris, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Cummings, N.J., Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D., Emmerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J., Grandi, G., Guiseppe, G., Guy, B.J., Haga, K., Halech, J., Harwood, C.R., Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H., Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R.P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M., Portetelle, D., Porwollik, S., Prescott, A.M., Presecan, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S.J., Serror, P., Shin, B.S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzenegger, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.F., Zumstein, E., Yoshikawa, H. and Danchin, A.

TITLE The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*

JOURNAL Nature 390 (6657), 249-256 (1997)

PUBMED 9384377

REFERENCE 2 (bases 1 to 1137)

CONSRM NCBI Genome Project

TITLE Direct Submission

JOURNAL Submitted (12-OCT-2001) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA

REFERENCE 3 (bases 1 to 1137)

AUTHORS Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.

TITLE Direct Submission

JOURNAL Submitted (18-NOV-1997) Institut Pasteur, Regulation de l'Expression Genetique, 28 rue du Docteur Roux, Paris Cedex 15 75724, France

COMMENT PROVISIONAL REFSEQ: This record has not yet been subject to final NCBI review. The reference sequence was derived from AL009126. On Jul 29, 2004 this sequence version replaced gi:16077068. COMPLETENESS: full length.

FEATURES

source 1..1137
/organism="Bacillus subtilis subsp. subtilis str. 168"
/mol_type="genomic DNA"
/strain="168"
/db_xref="taxon:224308"

gene 1..1137
/gene="dnaN"
/locus_tag="BSU00020"
/note="synonym: dnaG; dnaK"
/db_xref="GeneID:939970"

CDS 1..1137
/gene="dnaN"
/locus_tag="BSU00020"
/EC_number="2.7.7.7"
/function="DNA synthesis"
/note="binds the polymerase to DNA and acts as a sliding clamp"
/codon_start=1
/transl_table=11
/product="DNA polymerase III subunit beta"
/protein_id="NP_387883.1"
/db_xref="GI:16077070"
/db_xref="GOA:P05649"
/db_xref="UniProtKB/Swiss-Prot:P05649"
/db_xref="GeneID:939970"
/translation="MKFTIQKDLVESVQDVLKAVSSRTTIPILTGKIVASDDGVSE
TGSDSDISIESFIPKEEGDKELVTIEQPGSIVLQARFFSEIVKKLEPMATVEIEVQNY
LTIIRSGKAEFNLNGLDADEYPHLPQIEEHHAIQIPTDLLKNLIRQTVFAVSTSETRP
ILTVGNWKVEQSELLCTATDSHRLALRKAKLDIPEDRSYNVVLPKSLTELKILDDN
QELVDIVITETQVLFKAKNVLFSSRLLDGNYPDTTSLIPQDSKTEIIVNTRKFLQAI
D RASLLAREGRNNVVKLSAKPAESIEISSNSPEIGKVVEAIVADQIEGEELNISFSPKY
MLDALKVLEGAIEIRVSFTGAMRPFLIRTPNDETIVQLILPVRTY"

ORIGIN

```
1 atgaaattca cgattcaaaa agatcgtctt gttgaaagtg tccaagatgt attaaaagca
61 gtttcattcca gaaccacgat tcccattctg actgggtatta aaattgttgc atcagatgat
121 ggagtatoot ttacagggag tgactcagat atttotattg aatccttcac tccaaaagaa
181 gaaggagata aagaaatcgt cactattgaa cagcccggaa gcacgtcttt acaggctcgc
241 ttttttagtg aaattgtaaa aaaattgccc atggcaactg tagaaattga agtccaaaaat
301 cagtatttga cgattatccg ttctggtaaa gctgaattta atctaaacgg actggatgct
361 gatgaatata cgcacttgcc gcagattgaa gacatcatg cgattcagat cccaactgat
421 ttgttaaaaa atctaatac acaaacagta ttgcagtgt ccacctcaga aacacgcctt
481 atcttgacag gtgtaaactg gaaagtggag caaagtgaat tattatgcac tgcaacggat
541 agccaccgtc ttgcattaag aaaggcgaaa ctgatattc cagaagacag atcttataac
601 gtctgtgattc cgggaaaaag tttaactgaa ctacagcaaga ttttagatga caaccaggaa
661 ctgttagata tgcctatcac agaaacccaa gttctgttta aagcgaaaaa cgtctgttgc
721 ttctcacggc ttctggacgg gaattatcca gacacaacca gcttgattcc gcaagacagc
781 aaaacagaaa tcattgtgaa cacaaaagaa ttcttcagg ccattgatcg tgcattcttt
841 ttagctagag agggacgcaa caacgttgta aaactgtccg caaaacggc tgaatccatt
901 gaaattttctt ccaattcgcc agaaatcggg aaagtgtggg aagcaattgt tgcgatacaa
961 attgaagggtg aggaattaaa tatctctttt agtccaaaat atatgctgga tgcactaaag
```

```
1021 gtgcttgaag gaggagaaat agcgtaagc ttacaggeg caatgagac ttcttaatt
1081 cgcacgcga atgatgaaac gattgtacag cttatcttc ctgcagaac ctattaa
//
```

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Jul 24 2006 17:11:14



PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

Books

My NCBI
[Sign In] [Register]

Search Protein



for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenPept



Show 20



Send to

Range: from begin

to end

Features: ☒ CDD

Refresh

☐ 1: YP_089686. Reports DNA polymerase III...[gi:52783857]BLink, Conserved
Domains, Links

Comment Features Sequence

LOCUS YP_089686 378 aa linear BCT 03-DEC-2005

DEFINITION DNA polymerase III subunit beta [Bacillus licheniformis ATCC 14580].

ACCESSION YP_089686

VERSION YP_089686.1 GI:52783857

DBSOURCE REFSEQ: accession NC_006322.1

KEYWORDS .

SOURCE Bacillus licheniformis ATCC 14580 (DSM 13)

ORGANISM Bacillus licheniformis ATCC 14580
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

REFERENCE 1 (residues 1 to 378)

AUTHORS Veith,B., Herzberg,C., Steckel,S., Feesche,J., Maurer,K.H., Ehrenreich,P., Baeumer,S., Henne,A., Liesegang,H., Merkl,R., Ehrenreich,A. and Gottschalk,G.

TITLE The complete genome sequence of Bacillus licheniformis DSM13, an organism with great industrial potential

JOURNAL Unpublished

REFERENCE 2 (residues 1 to 378)

CONSRM NCBI Genome Project

TITLE Direct Submission

JOURNAL Submitted (28-SEP-2004) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA

REFERENCE 3 (residues 1 to 378)

AUTHORS Veith,B., Herzberg,C., Steckel,S., Feesche,J., Maurer,K.H., Ehrenreich,P., Baeumer,S., Henne,A., Liesegang,H., Merkl,R., Ehrenreich,A. and Gottschalk,G.

TITLE Direct Submission

JOURNAL Submitted (30-APR-2004) Institute of Microbiology and Genetics, Georg August University Goettingen, Goettingen Genomics Laboratory, Grisebachstr. 8, Goettingen D-37077, Germany

COMMENT VALIDATED REFSEQ: This record has undergone preliminary review of the sequence, but has not yet been subject to final review. The reference sequence was derived from AAU38993.
Method: conceptual translation.

FEATURES Location/Qualifiers

source 1..378
/organism="Bacillus licheniformis ATCC 14580"
/strain="DSM 13; ATCC 14580"
/db_xref="ATCC:14580"
/db_xref="taxon:279010"

Protein 1..378
/product="DNA polymerase III subunit beta"
/name="DNA polymerase III (beta subunit)"
/calculated_mol_wt=41862

Region 1..376
/region_name="beta_clamp"

```

      /note="Beta clamp domain; cd00140"
      /db_xref="CDD:29053"
Region  17..373
      /region_name="POL3Bc"
      /note="DNA polymerase III beta subunit; smart00480"
      /db_xref="CDD:47786"
CDS     1..378
      /gene="dnaN"
      /locus_tag="BLi00002"
      /coded_by="NC_006322.1:1828..2964"
      /note="binds the polymerase to DNA and acts as a sliding
      clamp"
      /transl_table=11
      /db_xref="GeneID:3098090"
ORIGIN
      1 mkftiqkdr1 vesvqdv1ka vssrttipil tgikivasde gvsitgsdsd isiesfipke
      61 dgdleivtie qpgsivlqar ffseivkk1p mstveievqn qyltiirsgk aefnlnglda
     121 seypllpqie ehhafqiptd l1kn1irqtv favstsetrp iltgvnwnvt ggelictatd
     181 shr1alrkak ldinedssyn vvipgksite lskilddhqe lvdivitetq vlfktnvlf
     241 fsrildgnyp dtnrlipqes ktnlivntke flqaidrasl laregrnnv klsaaanesi
     301 eissnspeig kvvetvnaeq iegedlkisf spkymidalk vlegedihvs ftgampfli
     361 rtpnddsivq lilpvrtv
//
```

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Jul 24 2006 17:32:14

EXHIBIT 2

```
#####
# Program: needle
# Rundate: Mon Jul 31 07:00:44 2006
# Align_format: srspair
# Report_file: /ebi/externserv/old-work/needle-20060731-07004323238664.output
#####
```

```
#=====
#
# Aligned_sequences: 2
# 1: SEQ173
# 2: G_kaustophilus
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1138
# Identity: 987/1138 (86.7%)
# Similarity: 987/1138 (86.7%)
# Gaps: 147/1138 (12.9%)
# Score: 5832.0
#
#=====
```

SEQ173	1	aattccgac	9
G_kaustophil	101	aactaacagcaaccgcctcgggagtaacgctgaccggaagcgattccgac	150
SEQ173	10	atttcaattgaatcggtttattccgcttgaaaaagaaggcaagttgctcgt	59
G_kaustophil	151	atttcaattgaatcggtttattccgcttgaaaaagaaggcaagttgctcgt	200
SEQ173	60	tgatgtgaaaagaccggggagcatcgactgcaggcgcgctttttctctg	109
G_kaustophil	201	tgacgtgaaaagaccggggagcatcgactgcaggcgcgctttttctctg	250
SEQ173	110	aaatcgtgaaaaaactgccgaacaaacggtggaaatcgaaacggaagac	159
G_kaustophil	251	aaatcgtgaaaaaactgccgaacaaacggtggaaatcgaaacggaagac	300
SEQ173	160	aactttttgacgatcatccgctcggggcaactcagaattccgcctcaatgg	209
G_kaustophil	301	aactttttgacgatcatccgctcggggcaactcagaattccgcctcaatgg	350
SEQ173	210	gctaaacgccgacgaatatccgcgcctgccgcaaattgaagaagaaaacg	259
G_kaustophil	351	gctaaacgccgacgaatatccgcgcctgccgcaaattgaagaagaaaacg	400
SEQ173	260	tgtttcaaataccggctgattttattgaaaaccgtgattcggcaaacgggtg	309
G_kaustophil	401	tgtttcaaataccggctgattttattgaaaaccgtgattcggcaaacgggtg	450
SEQ173	310	ttcgcggtttctacatcggaacgcgcgcccaatcttgacaggtgtcaactg	359
G_kaustophil	451	ttcgcggtttctacatcggaacgcgcgcccaatcttgacaggtgtcaactg	500
SEQ173	360	gaaagttgaacatggcgagcttgtctgcacagcgaccgacagtcacgct	409
G_kaustophil	501	gaaagttgaacatggcgagcttgtctgcacagcgaccgacagtcacgct	550


```
#####
# Program: needle
# Rundate: Mon Jul 31 07:01:59 2006
# Align_format: srspair
# Report_file: /ebi/externserv/old-work/needle-20060731-07015859433755.output
#####

#=====
#
# Aligned_sequences: 2
# 1: SEQ173
# 2: B_cereus
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1146
# Identity: 654/1146 (57.1%)
# Similarity: 654/1146 (57.1%)
# Gaps: 160/1146 (14.0%)
# Score: 3294.5
#
#=====

SEQ173      1                               aattccgac      9
                                     .|||.|||.
B_cereus    101 aggttgtagctacggaagaaggagttacattaacaggaagcgatgctgat 150

SEQ173      10 atttcaattgaatcgtttattccgcttgaaaaagaaggcaagttgctcgt      59
          |||||.|||||.|||||.|||.|||||.|||.|||||.|||.|||||.
B_cereus    151 atttcgattgaatctttatcccagttgaagaggatggaaaagaaatcgt 200

SEQ173      60 tgatgtgaaaagaccggggagcatcgactgcaggcgcgctttttctctg      109
          .|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
B_cereus    201 agaagtaaaacaatcaggaagtattgtttacaggcgaaatatttttagtg 250

SEQ173      110 aaatcgtgaaaaaactgccgcaacaaacggtggaaatcgaaacggaagac      159
          |||||.|||.|||||.|||.|||.|||.|||.|||.|||.|||.
B_cereus    251 aaattgtaaaaaaattacctaaagaaactgtagaaatttctgtggaaaat 300

SEQ173      160 aactttttgacgatcatccgctcggggcactcagaattccgcctcaatgg      209
          .|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
B_cereus    301 catttgatgacaaaaataacttctggaatcagagtttaatttaaattgg 350

SEQ173      210 gctaaacgccgacgaatatccgcgcctgcgcgcaaattgaagaagaaaacg      259
          .|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
B_cereus    351 tttagattctgcagaatatccgttggtaccacaaattgaagaacatcatg 400

SEQ173      260 tgtttcaaatcccgctgatttattgaaaaccgtgattcggcaaacggtg      309
          |||||.|||||.|||.|||||.|||.|||.|||.|||.|||.
B_cereus    401 tttttaaatacccaacggatttacttaacatatgatcagacaaactgta 450

SEQ173      310 ttgcgcgtttctacatcggaacgcgcgccaatcttgacaggtgtcaactg      359
          |||.|||||.|||.|||||.|||.|||.|||.|||.|||.
B_cereus    451 tttgcagtttccacttctgaaacaagaccgatcttgacaggtgtaaactg 500

SEQ173      360 gaaagttgaacatggcgagcttgtctgcacagcgaccgacagtcacatcgct      409
          |||.|||.|||.|||.|||.|||.|||.|||.|||.
B_cereus    501 gaaggtatataacagcgaaactaactgcattgtctacagatagccacaggt 550
```

SEQ173	410	tagccatgcgcaaaagtgaataattg-agtcg--gaaaatgaagtatcatac	456
B_cereus	551	tagctcttcgaaaagcaaaaatcgaggcattgcagatgaa-ttccaagc	599
SEQ173	457	aac-gtcgtcatccctggaaaaagtcttaatgagctcagcaaaattttgg	505
B_cereus	600	aaatgttggttatccagggaaaagcttaaatgaattaagcaaaattttag	649
SEQ173	506	atga--cggcaaccacccggtggacatcgatgacagccaatcaagtgc	553
B_cereus	650	atgagtcgtgaag--aaatggtagatatcggttattacggagtatcaagtat	697
SEQ173	554	tatttaaggccgagcacctctctctctcttcccggtgcttgacgggcaac	603
B_cereus	698	tattccgtacaaaacatttattattctctcaagattgttagaaggaaat	747
SEQ173	604	tatccggagacggcccgcttgattccaacagaaagcaaaacgaccatgat	653
B_cereus	748	tatcctgtatcaactcgcttaattccagcggagagtaaaacagatatttt	797
SEQ173	654	cgtcaatgcaaaaagagtttctgcaggcaatcgaccgagcgtctcttgcttg	703
B_cereus	798	tgtaaatacaaaaagaatttttacaagcaattgatcgatccctatttag	847
SEQ173	704	ctcgagaaggaaggaacaacgcttgtaaactgacgacgcttctcgagg-	752
B_cereus	848	caagagatggtcgtaataatgttgtaaaattatcaac-ttttagagcaggc	896
SEQ173	753	aatgctcgaaattttcttcgattttctcc-gagatcgggaaagt-gacggag	800
B_cereus	897	aatgctagaaattttcttcaaatccaccagaaatcgggaaagtagtagaag	946
SEQ173	801	cagctgcaaacggagctctctgaaggggaagagttgaacatttcgcttcag	850
B_cereus	947	aagtt-caatgtgaaaaagtagatggagaagagttaaaaatatcttttag	995
SEQ173	851	cgcgaaatatatgatggacgcgttgcgggcgcttgatggaacagacattt	900
B_cereus	996	tgcaaaatatatgatggatgcactaaaggcattagatagtactgaga-tt	1044
SEQ173	901	caaatcagcttcactggggccatgcggccgcttctgttgccgccgcttca	950
B_cereus	1045	aagattagctttactggagcaatgagaccattcttaattcgtacggtaaa	1094
SEQ173	951	-accgattcgatgcttcagctcattttgcccgttgagaacatat	992
B_cereus	1095	tgatgaatccattattcaattaattttacgggttcgtacttactaa	1140

```
#####
# Program: needle
# Rundate: Mon Jul 31 07:03:18 2006
# Align_format: srspair
# Report_file: /eh1/extserv/old-work/needle-20060731-07031637746308.output
#####

#=====
#
# Aligned_sequences: 2
# 1: SEQ173
# 2: B_thuringiensis
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1148
# Identity: 655/1148 (57.1%)
# Similarity: 655/1148 (57.1%)
# Gaps: 164/1148 (14.3%)
# Score: 3260.0
#
#=====

SEQ173      1                                aattccgac      9
                                     .|||.|||.
B_thuringiens 101 aagttgtcgtctacggaagaaggagttacattaacaggaagcgatgctgat 150

SEQ173      10 atttcaattgaatcgttttatccgcttgaaaaagaaggcaagttgctcgt      59
      |||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
B_thuringiens 151 atttcgattgaatctttatcccgagttgaagggatggaaaagaaattgt 200

SEQ173      60 tgatgtgaaaagaccggggagcatcgactgcaggcgcgctttttctctg      109
      .|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
B_thuringiens 201 agaagtaaaacaatcaggaagtattgttttacaggctaaatatttttagtg 250

SEQ173      110 aaatcgtgaaaaaactgccgcaacaaacgggtggaaa--tcgaaacggaag      157
      |||||.|||.|||||.|||||.|||.|||||.|||.|||||.|||.
B_thuringiens 251 aaattgtaaaaaattgccgaaagaaactgtgaaatttctc-tgtcgaaaa 299

SEQ173      158 acaactttttgacgatcatccgctcggggcactcagaattccgcctcaat      207
      .|||.|||.|||||.|||.|||.|||.|||.|||.|||.|||.
B_thuringiens 300 tcat-ttaatgacaaaaataacttctgggaatcagaatttaatttaaat 348

SEQ173      208 gggctaaaacgccgacgaatatccgcgcctgccgcaaattgaagaagaaaa      257
      |||.|||.|||.|||.|||||.|||.|||||.|||.|||||.|||.
B_thuringiens 349 ggttttagattctgcagaatatccattgttaccacaaattgaagaacatca 398

SEQ173      258 cgtgtttcaaatcccggtgattttattgaaaaccgtgattccggcaaacgg      307
      .|||.|||.|||.|||.|||||.|||.|||||.|||.|||||.|||.
B_thuringiens 399 tgtttttaagattccaacagatttacttaaacatatgatcagacaaactg 448

SEQ173      308 tgttcgcgctttctacatcggaacgcgcaccaatcttgacaggtgtcaac      357
      |||.|||.|||.|||.|||||.|||.|||||.|||.|||||.|||.
B_thuringiens 449 tatttgcagtcctcacttctgaaacaagaccaatcttgacaggtgttaaac 498

SEQ173      358 tggaaaagttgaacatggcgagcttgtctgcacagcgacgcagtcacgtcag      407
      |||||.|||.|||.|||.|||||.|||.|||||.|||.|||||.
B_thuringiens 499 tggaaaggtatataacagcgaaactaacttgattgtctacagatagtcacag 548
```

SEQ173	408	cttagccatgcgcgcaagtgaaaaattga--gtc-ggaaaaatgaagtatcat	454
B_thuringiens	549	gttagctcttcgaaaaagcaaaaattgaaggtattgtagatgaa-ttccag	597
SEQ173	455	acaac-gtcgtcatccctggaaaaagtcttaatgagctcagcaaaatttt	503
B_thuringiens	598	gcaaatgttggtatttcggggaaaaagcttaaatgaattaagcaaaattct	647
SEQ173	504	ggatga--cggcaaccaccgggtggacatcgctcatgacagccaatcaagt	551
B_thuringiens	648	agatgagctcgaag--aaatggtagatatcggtattacggagtatcaagt	695
SEQ173	552	gctatttaaggccgagcaccttctctcttttccggctgcttgacggca	601
B_thuringiens	696	attatccgtacaaaaacatttattattcttctcaagattgtagaaggaa	745
SEQ173	602	actatccggagacggcccgcttgattccaacagaaaacgacccatg	651
B_thuringiens	746	attatcctgatacaactcgatttaattcctgcagagagtaaacagatatt	795
SEQ173	652	atcgctcaatgcaaaagagtttctgcaggcaatcgaccgagcgctccttgc	701
B_thuringiens	796	ttgtaaaatacaaaaagaattttacaagcaattgatcgctcatctctgtt	845
SEQ173	702	tgctcgagaaggaaggaacaacggttgtaaactgacgacgcttcctggag	751
B_thuringiens	846	agctagagatggtcgtaataatggttgaaaattatcaac-tttagagcag	894
SEQ173	752	g-aatgctcgaaaatttcttcgattttctcc-gagatcgggaaagt-gacgg	798
B_thuringiens	895	gcaatgctagaaaatttcttcgaattcaccagaaatcgggaaagttagtaga	944
SEQ173	799	agcagctgcaaacggagctctcttgaaggggaagagttgaacatttcgttc	848
B_thuringiens	945	agaagtt-caatgtgaaaaagtagatggagaagagttaaaaatatctttt	993
SEQ173	849	agcgcgaaatatatgatggacggttgcgggcgcttgatggaacagacat	898
B_thuringiens	994	agtgcaaaatatatgatggatgcactaaaggcattagatagtactga-aa	1042
SEQ173	899	ttcaaatcagcttcactggggccatgcggccgttctctgttgccgccgctt	948
B_thuringiens	1043	ttaagattagctttactggagcaatgagaccattcttaattcgtacggta	1092
SEQ173	949	ca-accgattcgatgcttcagctcattttgcccgtgagaacatat	992
B_thuringiens	1093	aatgatgaatccattattcaattattttaccggttcgtactactaa	1140

```
#####
# Program: needle
# Rundate: Mon Jul 31 07:04:20 2006
# Align_format: srspair
# Report_file: /ebi/extserv/old-work/needle-20060731-07041940078774.output
#####
```

```
#####
#
# Aligned_sequences: 2
# 1: SEQ173
# 2: B_weihenstephanensis
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1149
# Identity: 636/1149 (55.4%)
# Similarity: 636/1149 (55.4%)
# Gaps: 160/1149 (13.9%)
# Score: 3242.0
#
#####
```

SEQ173	1	aattccgac	9
B_weihensteph	101	aagttgtagcaactgaagaaggagttactttaacaggtagtgatgcagac	150
SEQ173	10	atttcaattgaatcgtttattccgcttgaaaaagaaggcaagttgctcgt	59
B_weihensteph	151	atctctatcgaaatcatttatcccagtcgaagatgctggaaaagaaattgt	200
SEQ173	60	tgatgtgaaaagaccggggagcagtcgtaactgcaggcgcgctttttctctg	109
B_weihensteph	201	ggaaattgaacaatcaggaagtattattttacaagcaaatattttagtg	250
SEQ173	110	aaatcgtgaaaaaactgcgcgaacaaacgggtggaaatcgaaacggaagac	159
B_weihensteph	251	agatcgtaaaaaaattacctaagaaactgtcgaaatttctgtggaaat	300
SEQ173	160	aactttttgacgatcatccgctcggggcactcagaattccgcctcaatgg	209
B_weihensteph	301	cattttatgacaaaaataaaatctggaaaatcagagttcaacttaaattg	350
SEQ173	210	gctaaacgcgcgaagaatatacgcgcctgcgcgaatgaagaagaaaacg	259
B_weihensteph	351	tttagatgctgctgaatatccgttattaccacaaatcgaagaacatcatg	400
SEQ173	260	tgtttcaaatcccggtgatttattgaaaaccgtgattcggcaaacggtg	309
B_weihensteph	401	tatttaagattccaacagatttactaaagcacatgattcgtcaaacagta	450
SEQ173	310	ttcgccgtttctacatcggaacgcgcaccaatcttgacaggtgtcaactg	359
B_weihensteph	451	tttgagtttcaagttctgaaacaagacccgatcttgacaggtgtaaaactg	500

SEQ173	360	gaaagttgaacatggcgagcttgtctgcacagcgaccgacagtcacgcgt	409
B_weihensteph	501	gaaggtatataaacagcgagctaacttgattgcaacagatagtcacagac	550
SEQ173	410	tagccatgcgcgaaagtgaattga---gtcggaaaatgaag-----tat	451
B_weihensteph	551	tagcacttcgtaaaagcaaaaatcgaagggtataatattgccgatgaattt	600
SEQ173	452	catac-aacgctcgtcacccctggaaaaagtcttaatgagctcagcaaaat	500
B_weihensteph	601	caagcgaatgtcggtattctcgttggaagagcttaagtgaattaagtaaaat	650
SEQ173	501	tttgatga--cggcaaccaccgggtggacatcgatcacagccaatca	548
B_weihensteph	651	tctagatgaatctgaag--aaatggtagatatcggtattacggagtatca	698
SEQ173	549	agtgctattttaaggccgagcaccttctctcttttccggctgcttgacg	598
B_weihensteph	699	agtattattccgtacaaaacatttattattcttctcaagattattagaag	748
SEQ173	599	gcaactatccggagacggcccgcttgattccaacagaaagcaaaacgacc	648
B_weihensteph	749	gaaattatccagatacaacgcgtttaattccagctgaaagtaaaactgat	798
SEQ173	649	atgatcgtaaatgcaaaagagtttctgcaggcaatcgaccgagcgctcctt	698
B_weihensteph	799	atatttgtaaatacaaaagaatttttacaagcgattgatcgctgcgtcgtt	848
SEQ173	699	gcttgctcgagaaggaaggaacaacgcttggtgaaactgacgacgcttcctg	748
B_weihensteph	849	gttagcaagagatggctgtaacaacgcttggttaaattatcgacattagaac	898
SEQ173	749	gaggaatgctcgaaattttcttcgatttctcc-gagatcgggaaagtgcg	797
B_weihensteph	899	aacagatgtagagatttcttcgaatgcaccagaaatcggaagtagta	948
SEQ173	798	gagcagctgcaaacggagtcctcttgaagggaagagttgaacatttcgtt	847
B_weihensteph	949	gaagaggttcaatgtgaaaatgtagatggagaagaattaaaaatatcttt	998
SEQ173	848	cagcgcgaaatatatgatggacgcgttgccggcgcttgatggaacagaca	897
B_weihensteph	999	tagtgcaaaatatatgatggacgcattaaaggcttttagacagtacaga-a	1047
SEQ173	898	tttcaaatcagcttcaactggggccatgcggccgcttcctggtgcgccgct	947
B_weihensteph	1048	attaaagttagttttactggagcgatgagaccgtttttaattgcacagt	1097
SEQ173	948	tca-accgattcgatgcttcagctcattttgcoggtgagaacatat	992
B_weihensteph	1098	aaatgatgattccattatccaattaattttaccagtcogtaactactaa	1146

SEQ173	408	cttagccatgcgcaaaagtgaattgagtcggaaaatgaagtatcataca	457
B_subtilis	549	tcttgcatgaagaaaggcgaaacttgatattccagaagacagatcttata	598
SEQ173	458	acgtcgtcatccctggaaaaagtcttaatgagctcagcaaaatcttgat	507
B_subtilis	599	acgtcgtgattccgggaaaaagtcttaactgaactcagcaagattttagat	648
SEQ173	508	gacggcaaccaccgggtggacatcgatgacagccaatcaagtgtctatt	557
B_subtilis	649	gacaaccagggaacttgtagatcgatcatcacagaaacccaagtctgtt	698
SEQ173	558	taaggccgagcaccttctc-ttcttttcccggtgcttgacggcaactat	606
B_subtilis	699	taaag-cgaaaaacgtcttcttctcaccggtcttgacgggaattat	747
SEQ173	607	cgggagacggcccgcttgattccaacagaaagcaaacgaccatgatcgt	656
B_subtilis	748	ccagacacaaccagcctgattccgcaagacagcaaacagaaatcattgt	797
SEQ173	657	caatgcaaaagagtttctgcaggcaatcgaccgagcgtcttctgtctc	706
B_subtilis	798	gaacacaaaagaattccttcaggccattgatcgtgcattctcttttagcta	847
SEQ173	707	gagaagggaaggacaacgcttgtaaactgacgacgcttcttgagggaatg	756
B_subtilis	848	gagagggacgcaacaacgcttgtaaactgtccgcaaaacc-ggctgaatc	896
SEQ173	757	c-tcgaaatttcttcgattttctcc-gagatcgggaaag-tgacggagcag	803
B_subtilis	897	cattgaaatttcttccaattcgccagaaatcggtaaagtgtggaagcaa	946
SEQ173	804	ctgcaaacggagttctcttgaagggaagagttgaacatttcttcagcgc	853
B_subtilis	947	ttg-ttgcggtacaaattgaaggtgaggaattaaatatctcttttagtcc	995
SEQ173	854	gaaatatatgatggacgcttgcgggcgcttgatggaacagacatttcaa	903
B_subtilis	996	aaaatatatgctggatgcactaaaggtgcttgaaggagcaga-aatacgc	1044
SEQ173	904	atcagcttcaactggggccatgcggccgttctgttgccgccgcttcaacc	953
B_subtilis	1045	gtaagctttacaggcgcaatgagacctttcttaattcgcacgc--cgaat	1092
SEQ173	954	gat---tcgatgcttcagctcattttgcgggtgagaacatat	992
B_subtilis	1093	gatgaaacgattgtacagcttctcctctgtcagaacctattaa	1137

```

#####
# Program: needle
# Rundate: Mon Jul 31 07:09:46 2006
# Align_format: srspair
# Report_file: /ebi/extserv/old-work/needle-20060731-07094578887427.output
#####

#=====
#
# Aligned_sequences: 2
# 1: SEQ173
# 2: B_licheniformis
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1161
# Identity: 704/1161 (60.6%)
# Similarity: 704/1161 (60.6%)
# Gaps: 193/1161 (16.6%)
# Score: 3657.5
#
#=====

SEQ173      1                                aattccgac      9
                                     .|||||.
B_licheniform 101 aaatcgtggcctctgatgaaggggtctctctgacaggcagcgattccgat 150

SEQ173      10 atttcaattgaatcgtttatccgcttgaaaaagaaggcaagtt---gct      56
      |||||.|||||.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
B_licheniform 151 atttcgattgaatcgtttatcccg---aaagaagacggcgatttagagat 197

SEQ173      57 cgttg-atgtgaaaagaccggggagcatcgtaactgcaggcgcgctttttc      105
      |||. |.|||||.|||. |.|||||.|||. |.|||||.|||. |.|||||.
B_licheniform 198 cgtgacaattgaacagccc-ggcagcattgtgcttcaagcccgcttttttc 246

SEQ173      106 tctgaaatcgtgaaaaaactgcgcgaacaaacggtggaaatcgaaacgga      155
      ..|||||.|||.|||||.|||||. ....|||.|||||.|||||. ....|
B_licheniform 247 agtgaaattgtcaaaaagctgccgatgtcaacagtggaaatcgaggttca 296

SEQ173      156 agacaactttttgacgatcatccgctcggggcactcagaattccgcctca      205
      |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |
B_licheniform 297 aaatcaatacttaacgatcatccgctcgggcaaagcagagttaacttaa 346

SEQ173      206 atgggctaaacgccgacgaatatccgcgcctgccgcaaattgaagaagaa      255
      |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |
B_licheniform 347 acggtttggatgcaagcgaatatccgcttttgccgcaaattgaagagcat 396

SEQ173      256 aacgtgtttcaaatcccggtgattttattgaaaaccgtgattcggaac      305
      .|||.|||||.|||||.|||||. |. |. |. |. |. |. |. |. |. |
B_licheniform 397 cacgcttttcaaatccgacgatctgctgaaaaacctgatccgcaaac 446

SEQ173      306 ggtgttcgcgctttctacatcggaacgcgcccaatcttgacaggtgtca      355
      .|||. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |
B_licheniform 447 cgtttttgcagtgtccacctcagaaacacgcccaatcttgacaggtgtaa 496

SEQ173      356 actggaaagttgaacatggc---gagcttgtctgcacagcgacagcagc      402
      |||||. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |
B_licheniform 497 actggaatgt--cac-tggcgggtgaattaatatgcactgcaacggatagt 543

```

SEQ173	403	catcgcttagccatgcgcaaaagtgaattgag-tcgga--aatgaagt	449
B_licheniform	544	catc----gtcttgcgctaag-gaaagctaagctogacattaacgaaga	587
SEQ173	450	---atcatacaacgtcgtcacccctggaaaaagtottaatgagctcagca	496
B_licheniform	588	cagttcatacaatgtcgtcacccaggaaaaagcttaaccgagctcagca	637
SEQ173	497	aaattttggatgacggcaacc--acccggtggacatcgtcacgagcca	544
B_licheniform	638	aaatccttgatgac--catcaggagcttgttgatattgtgattaccgaaa	685
SEQ173	545	atcaagtgtctatttaaggccgagcaaccttctcttcttctccggctgctt	594
B_licheniform	686	cacaagtgtgtttaaaacaaaaacgttctgttttctccaggcttctt	735
SEQ173	595	gacggcaactatccggagacggcccgcttgattccaacagaaagcaaac	644
B_licheniform	736	gacggaaactatccggatacgaaccgctgattcctcaggaaagcaaac	785
SEQ173	645	gaccatgatcgtaaatgcaaaagagtttctgcaggcaatcgaccgagcgt	694
B_licheniform	786	gaacttgattgtcaataactaaggaaattctctccaggcgatcgacagggtt	835
SEQ173	695	ccttgctt--gctcgagaagggaaggaacaacgttgtgaaactgacgacg	741
B_licheniform	836	c---gcttttggcgagagaaggacgaacaacgttgtgaaacttctc--cg	880
SEQ173	742	cttc--ctggaggaatgctcgaaatttcttcgatttctcc-gagatcggg	788
B_licheniform	881	cagccgccaatgagtcgatcgaaatttcttcaaactctccggaatcgga	930
SEQ173	789	aaag-tgacggagcagctgcaaacggagttctcttgaagggaagagttga	837
B_licheniform	931	aagggtgttgaaacgg-tgaatgccgagcagatcgaagggaagacttaa	979
SEQ173	838	acatttcggttcagcgcgaaatatatgatggacgcgttgccggcgcttgat	887
B_licheniform	980	agatatcctttagtcggaatatatgctggatgcccttaaagtctctgaa	1029
SEQ173	888	ggaacagacatttcaaatcagcttcaactggggccatgcggccggttctgt	937
B_licheniform	1030	ggagaggaca-ttcatgtaagcttcacaggcgctatgaggcctttctga	1078
SEQ173	938	tgcgcccgttcaa--ccgattcgatgcttcagctcattttgcgggtga	984
B_licheniform	1079	tccgtacgc--cgaatgacgattcgatcgccaattaattcttctgtcc	1126
SEQ173	985	gaacatat	992
B_licheniform	1127	ggacgtattaa	1137

.....

CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ174 334 aa

Sequence 2: G_kaustophilus 378 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 97

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05543637.dnd]

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:6996

Alignment Score 1906

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05543637.aln]

```

SEQ174 -----NSDISIIIESFIPL 13
G_kaustophilus MNISIDREALAKSVQDVMKAVSTRRTTIPILTGIKLTATASGVTLTGSDSDIS-IESFIPL 59
                      :****

SEQ174 EKEGKLLVDVKRPGSIVLQARFFSEIVKKLPQQTVEIETEDNFLTIIIRSGHSEFRLNGLN 73
G_kaustophilus EKEGKLLVDVKRPGSIVLQARFFSEIVKKLPQQTVEIETEDNFLTIIIRSGHSEFRLNGLN 119
                      *****

SEQ174 ADEYPRLPQIEEENVFQIPADLLKTIVIRQTVFAVSTSETRPILTGVNWKVEHGELVCTAT 133
G_kaustophilus ADEYPRLPQIEEENVFQIPADLLKTIVIRQTVFAVSTSETRPILTGVNWKVEHGELVCTAT 179
                      *****

SEQ174 DSHRLAMRKVKIIESENEVSYNVVIPGKSLNELSKIILDDGNHPVDIVMTANQVLFKAEH 193
G_kaustophilus DSHRLAMRKVK-IESENEVSYNVVIPGKSLNELSK-ILDDGNHPVDIVMTANQVLFKAEH 237
                      *****

SEQ174 LLFFSRLLDGNYPETARLIPTESKTTMIVNAKEFLQAI DRASLLAREGRNNVVKLTTLPG 253
G_kaustophilus LLFFSRLLDGNYPETARLIPTESKTTMIVNTKEFLQAI DRASLLAREGRNNVVKLTTLPG 297
                      *****

SEQ174 GMLEISSISPEIGKVTEQLQTESLEGEELNISFS AKYMMDALRALDGTDIQISFTGAMRP 313
G_kaustophilus GMLEISSISPEIGKVTEQLQTESLEGEELNISFS AKYMMDALRALDGTDIQISFTGAMRP 357
                      *****

SEQ174 FLLRPLHTDSMLQLILPVRTY 334
G_kaustophilus FLLRPLHTDSMLQLILPVRTY 378
                      *****

```

CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ174 334 aa

Sequence 2: B_cereus 379 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 70

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05564381.dnd]

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:6158

Alignment Score 1389

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05564381.aln]

```

SEQ174      -----NSDISI-IESFIPLE 14
B_cereus    MRFTIQKDYLVRSVQDVMKAVSSRTTIPILTGIVVATEEGVTLTGSADADISIESFIPVE 60
              .** . * *****;*

SEQ174      KEGKLLVDVKRPGSIVLQARFFSEIVKKLPQQTVEIETEDNFLTIIIRSGHSEFRLNGLNA 74
B_cereus    EDGKEIVEVKQSGSIVLQAKYFSEIVKKLPKETVEISVENHLMTKITSGKSEFNLNGLDS 120
              :*** :*:***:*****:*****:*****:*****:*****:*****:*****:
              :***:***:*****:*****:*****:*****:*****:*****:*****:

SEQ174      DEYPRLPQIEEENVFQIPADLLKTVIRQTVFAVSTSETRPILTGWNWKVEHGEVLVCTATD 134
B_cereus    AEYPLLPQIEEHHVFKEIPTDLLKHMIRQTVFAVSTSETRPILTGWNWKVYNSELTCIATD 180
              *** *****:***:***:*****:*****:*****:*****:*****:*****:
              :***:***:*****:*****:*****:*****:*****:*****:*****:

SEQ174      SHRLAMRKVKIIESENEVSYNVVIPGKSLNELSKIILDDGNHPVDIVMTANQVLFKAHL 194
B_cereus    SHRLALRKAKIEGIADEFQANVVIPGKSLNELSK-ILDESEEMVDIVITEYQVLFRTKHL 239
              *****:***:***:*****:*****:*****:*****:*****:*****:
              :***:***:*****:*****:*****:*****:*****:*****:*****:

SEQ174      LFFSRLLDGNYFPETARLIPTESKTTMIVNAKEFLQAIDRASLLAREGRNNVVKLTTLPGG 254
B_cereus    LFFSRLLEGNYPDTRRLIPAESKTDIFVNTKEFLQAIDRASLLARDGRNNVVKLSTLEQA 299
              *****:***:***:*****:*****:*****:*****:*****:*****:
              :***:*****:*****:*****:*****:*****:*****:*****:*****:

SEQ174      MLEISSISPEIGKVTEQLQTESLEGEELNISFSKYMMDALRALDGTDIQISFTGAMRPF 314
B_cereus    MLEISSNSPEIGKVVEEVQCEKVDGEELKISFSKYMMDALKALDSTEIKISFTGAMRPF 359
              ***** *****:***:***:*****:*****:*****:*****:*****:*****:
              :***:*****:*****:*****:*****:*****:*****:*****:*****:

SEQ174      LLRPLHTDSMLQLILPVRTY 334
B_cereus    LIRTVNDESIIQLILPVRTY 379
              ***:***:*****:*****:*****:*****:*****:*****:*****:

```


CLUSTAL W (1.83) Multiple Sequence Alignments

```

Sequence format is Pearson
Sequence 1: SEQ174                334 aa
Sequence 2: B_thuringsiensis     379 aa
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned. Score: 70
Guide tree          file created:  [/ebi/extserv/clustalw-work/interactive/clustalw-
20060731-06031038.dnd]
Start of Multiple Alignment
There are 1 groups
Aligning...
Group 1: Sequences: 2             Score:6155
Alignment Score 1387
CLUSTAL-Alignment file created  [/ebi/extserv/clustalw-work/interactive/clustalw-
20060731-06031038.aln]

```

```

SEQ174                                -NSDISI-IESFIPLE 14
B__thuringiensis                     MRFSIQKDYLVRSVQDVMKAVSSRTTIPILTGIKVVATEEGVTLTGSADISIESFIPVE 60
                                         .** .* *****:
                                         .

SEQ174                                KEGKLLVDVKRPGSGIVLQARFFSEIVKKLPQQTVEIETEDNFLTIIRSGHSEFRLNGLNA 74
B__thuringiensis                     EDGKEIVEVKQSGSIVLQAKYFSGEIVKKLPKETVEISVENHMLTKITSGKSEFNLNGLDS 120
                                         ::** :*:**:*****:*****:*****:*****:*****:*****:
                                         :

SEQ174                                DEYPRLPQIEEENVFQIPADLLKTVIRQTVFVAVSTSETRPILTGVNWVKVEHGEIVCTATD 134
B__thuringiensis                     AEYPLLPQIEEHHVFKIPTDLLKHMIRQTVFVAVSTSETRPILTGVNWKVYNSELTCIATD 180
                                         *** *****:*****:*****:*****:*****:*****:*****:
                                         :.*** *

SEQ174                                SHRLAMRKVKIIESENEVSYNVVIPGKSLNELSKIILDDGNHPVDIVMTANQVLFKAEHL 194
B__thuringiensis                     SHRLALRKAKIEGIVDEFQANVVIIPGKSLNELSK-ILDESEEMVDIVITEYQVLFRTKHL 239
                                         *****:*.** :*. ***** ***** :. *****: *****:
                                         :

SEQ174                                LFFSRLLDGNYPETARLIPTESKTTMIVNAKEFLQAIADRASLLAREGRNNVVKLTTLPGG 254
B__thuringiensis                     LFFSRLLLEGNYPDTRRLIPAESKTDIFVNTKEFLQAIADRASLLARDGRNNVVKLTSTLEQA 299
                                         *****:*****:*****:***** :*:*****:*****:*****:*****:
                                         :

SEQ174                                MLEISSISPEIGKVTEQLQTESLEGEELNISFSAKYMMDALRALDGTDIQISFTGAMRPF 314
B__thuringiensis                     MLEISSNSPEIGKVVEEVQCEKVDGEELKISFSAKYMMDALKALDSTEIKISFTGAMRPF 359
                                         ***** *****:*.** *:*:*****:*****:*****:*****:
                                         :

SEQ174                                LLRPLHTDSMLQLILPVRTY 334
B__thuringiensis                     LLRTVNDESIIQLILPVRTY 379
                                         *:*.**: :*:*****
                                         :

```

CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ174 334 aa

Sequence 2: *B. weihenstephanensis* 381 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 69

Guide tree file created: [/ebi/extserv/clusterw-work/interactive/clusterw-20060731-06054661.dnd]

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:6123

Alignment Score 1374

CLUSTAL-Alignment file created [/ebi/extern/clusterw-work/interactive/clusterw-20060731-06054661.aln]

```

SEQ174
E_weihenstephanensis
-----NSDISI-IESFIPLE 14
MRFTIQKDYLVRSVQDVMKAVSSRTTIPILTGIKVVATEEGVTLTGSDADISIESFIPVE 60
                      **. * *****.

```

SEQ174 KEGKLLVDVKRPGSIVLQARFFSEIVKKLPQQTVEIETEDNFLTIIRSGHSEFRLNGLNA 74
B_welhenstephanensis DAGKEIVEIEQSGSIILQAKYFSEIVKKLPKEIVSEIENHFMTKIKSGKSEFNLNGLDA 120
. ** :*::::***:***:*****:*****.:*:*:*:*:*:*:*:*:*:*:

SEQ174
B_weihenstephanensis
DEYPRLPQIEEENVFQIPADLLKTVIRQTVFAVSTSETRPILTGVNWKVEHGLVCTATD 134
AEYPLLPQIEEHHVFQIPTDLLKHMIRQTVFAVSSSETRPILTGVNWKVYNSELTCIATD 180
*** ***,*:*:*** :*****:*****:*** **

```

SEQ174          SHRLAMRKVKVIES--ENEVSYNVVI PGKSLNELSKI ILDDGNHPVDIVMTANQVLFKAE 192
B_weihenstephanensis SHRLALRKAKIEGYNIADEFQANVVI PGKSLSEL SK-ILDESEEMVDIVITEYQVLFRTK 239
*****:*.*. :*..*****.*.*** ***:..*:*:*:*

```

```

SEQ174      HLLFFSRLLDGNYPETARLIPTESKTTMIVNAKEFLQAI DRASLLAREGRNNVVKLTTLF 252
B_weihenstephanensis HLLFFSRLLLEGNYPDTRLIPAESKTDIFVNTKEFLQAI DRASLLARDGRNNVVKLSTLE 299
*****:*::*:***:**** :*:**:*****:*****:**

```

SEQ174 GGMLEISSISPEIGKVTEQLQTESLEGEELNISFSAKYMMDALRALDGTDIQISFTGAMR 312
B_weihenstephanensis QQMLEISSNAPEIGKVVEEVQCENVDGEELKISFSAKYMMALKALDSTEIKVSFTGAMR 359
***** :*****.*:* *.::****:*****:***.*::*****

```

SEQ174          PELLRLPLHTDSMLQLILPVRTY 334
E_weihenstephanensis PELLIRTVNDDSIQLILPVRTY 381
***:*::: **::*****

```

Sequence format is Pearson

Sequence 2: B subtilis 378 aa

Aligning...

```
Guide tree      file created:  [/ebi/extserv/clustalw-
```

Start of Multiple Alignment

Aligning...

Alignment Score 1379

```

SEQ174      -----NSDISIIIESFIPL 13
B_subtilis  MKFTIQKDRLVESVQDVLKAVSSRTTPIILTGIKIVASDDGVSETGSDSDIS-IESFIPK 59
                                     :**** *

```

SEQ174 ADEYPRLPQIEEENVFQIPADLLKTVIROTTFVAVSTSETRPILTGWNWKEHGEVLCTAT 133
B_subtilis ADEYPLHPQIEEHHAIQIPTDLLKNLIRQTFFVAVSTSETRPILTGWNWKVEQSSELLCTAT 179

*****::*:***.:*****:.....**:*:

```

SEQ174      LLFFSRLLDGNYPETARLIPTESKTTMIVNAKEFLQAIDRASLLAREGRNNVVKLTTPLG 253
B_subtilis VLFFSRLLDGNYPDTTSLIPQDSKTEIIVNTKEFLQAIDRASLLAREGRNNVVKLSAKPA 297
:*****: : * * * : * * * : * * * :*****: : *

```

```

SEQ174          FLLRPLHTDSMLQLILPVRTY 334
B_subtilis     FLIRTPNDETIVQLILPVRTY 378
                **:* : : : *****

```

Sequence format is Pearson

```

SEQ174
B_licheniformis
-----NSDISIIIESFIPL 13
MKFTIQKDRLVESVDVLKAVSSRTTIPILTGIKIVASDEGVSLTGSDDIS-IESFIKP 59
:**** *****

SEQ174
B_licheniformis
EKEGKLLVDVKRPGSIVLQARFFSEIVKKLPQQTVEIETEDNFLTIIIRSGHSEFRLNGLN 73
EDGDLEIVTIEQPGSIVLQARFFSEIVKKLPMSTVEIEVQONQYLTIIIRSGKAEFNLNGLD 119
*. . :* ::*****.*****.:::*****:*.****:

SEQ174
B_licheniformis
ADEYPRLPQIEEENVFQIPADLLKTVIRQTVFAVSTSETRPILTVGNWVKVEHGELVCTAT 133
ASEYPLLPQIEEHAFQIPTDLLKNLIRQTVFAVSTSETRPILTVGNWNVVTGGELICTAT 179
*.*** *****.:.****:***.:*****:*****:*.***:***

SEQ174
B_licheniformis
DSHRLAMRKVKIIESENEVSYNVVIPIGKSLNELSKIIILDDGNHPVDIVMTANQVLFKAEH 193
DSHRLALRKAKLDINE-DSSYNVVIPIGKSLTELSK-ILDDHQELVDIVITETQVLFKTKN 237
*****:*.*: .*: *****.***** ***** :. *****:*.*****::

SEQ174
B_licheniformis
LLFFSRLLDGNYPETARLIPTESKTTMIVNAKEFLQAI DRASLLAREGRNVVKLTTLTPG 253
VLFFSRLLDGNYPDTNRLIPQESKTNLIVNTEFLQAI DRASLLAREGRNVVKLSAAAN 297
:*****:*. *****:***:*****:*****:*****:*.

SEQ174
B_licheniformis
GMLEISSISPEIGKVTEQLQTESLEGEELNISFSKAYMMDALRALDGTDIQISTGAMRP 313
ESIEISSNSPEIGKVETVNAEQIEGEDLKISFSPKYMMLDALKVLEGEDIHVSTGAMRP 357
:*** *****.* ::*:***:***:***.***:***:*.*:*****

SEQ174
B_licheniformis
FLLRPLHTDSMLQLILPVRTY 374
FLIRTPNDDSIQVQLILPVRTY 378
***.*:***.*****

```